

# The role of a work identifier in the publishing process

---

*The Digital Supply Chain: today and tomorrow*

London: 4 December 2008

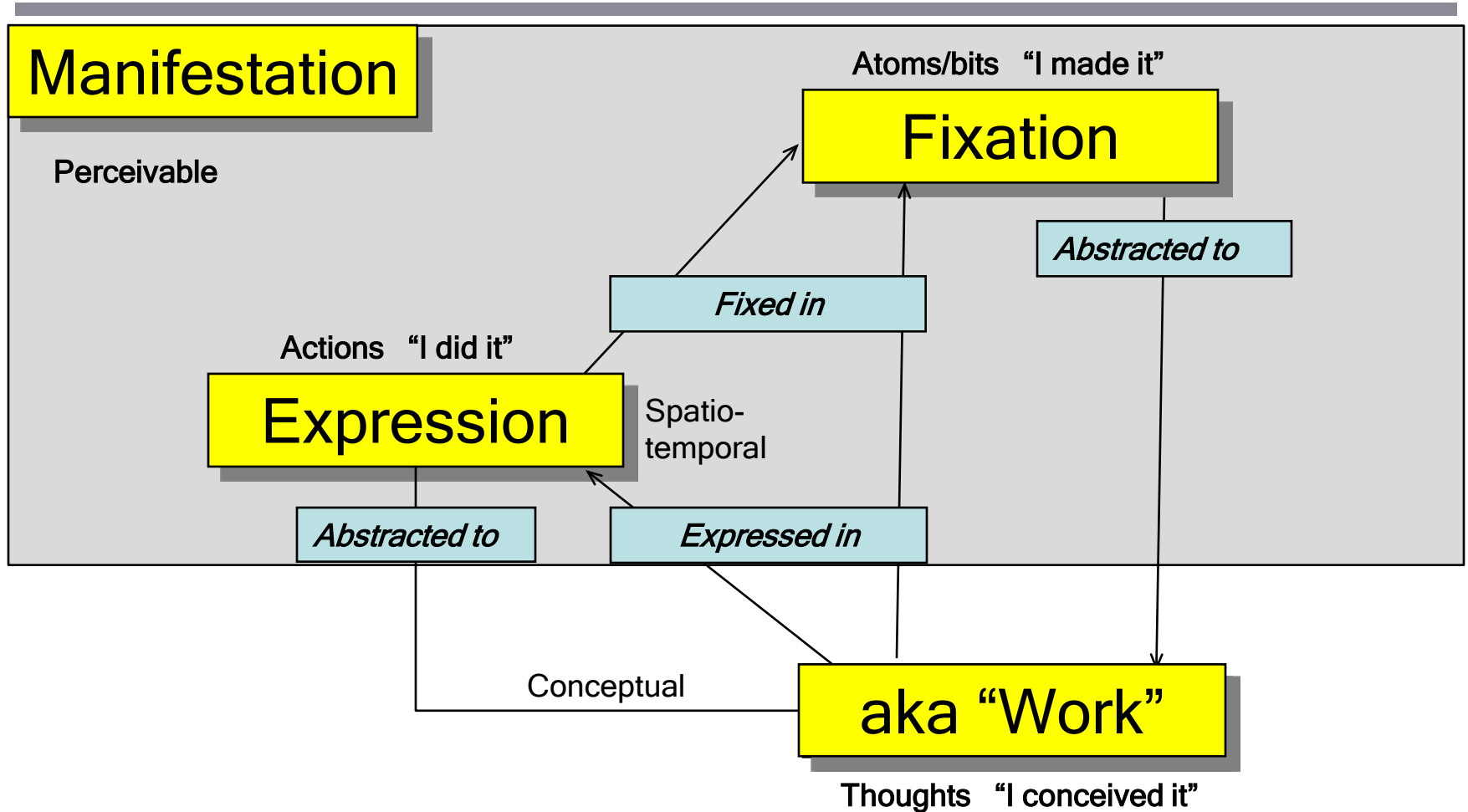
**Mark Bide, Rightscom Limited**

---

Beginning at the beginning

# WHAT IS "A WORK"?

# The primary types of creation: the <index> model of making

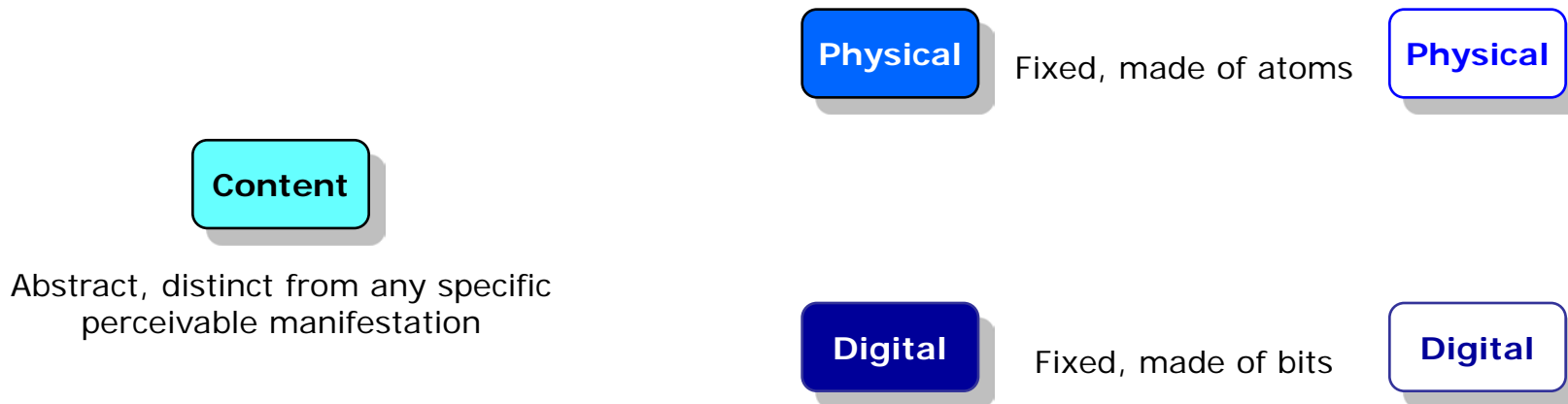
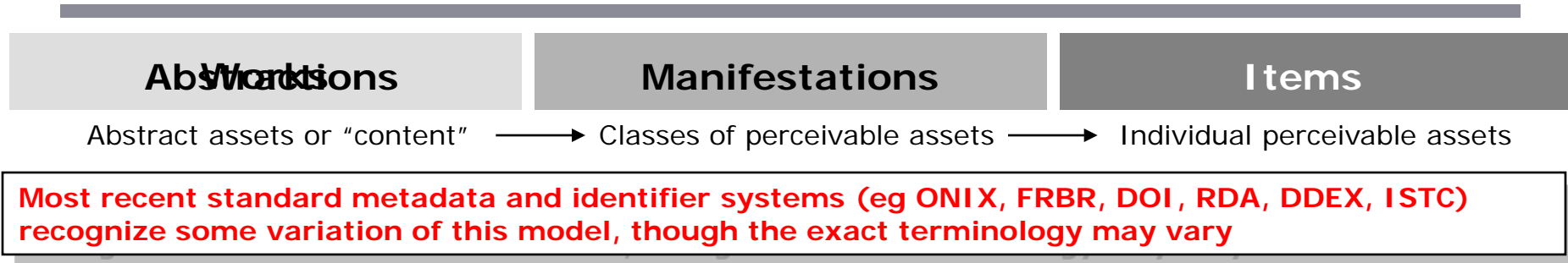


---

An example of the requirement for work identification in the publishing process

# **DIGITAL ASSET MANAGEMENT**

# Classes of publishers' assets



**Manifestations** are classes of physical or digital assets with common identifiers and identical attributes. For example: the class of all books with the same ISBN is viewed as a single **Manifestation**. An individual copy of the book is an **Item**.

# Classes of publishers' assets

## Abstractions

Abstract assets or "content"

eg

the words of a book or article  
the image in a photograph  
the figures and layout of a table  
a logo  
the design of a chair  
a graph

**Content**

or the complete contents of

a volume of a journal  
a magazine  
a book of photographs and maps  
a series of books  
a website  
an academic course pack  
a tv soap opera

## Manifestations

Classes of perceivable assets

**Physical**

eg, a class of  
printed book  
music CD  
leaflet  
sculpture

**Digital**

eg, a class of  
.pdf of an article  
.mp3 of a sound recording  
.exe of a program  
.gif of a photograph

## Items

Individual perceivable assets

**Physical**

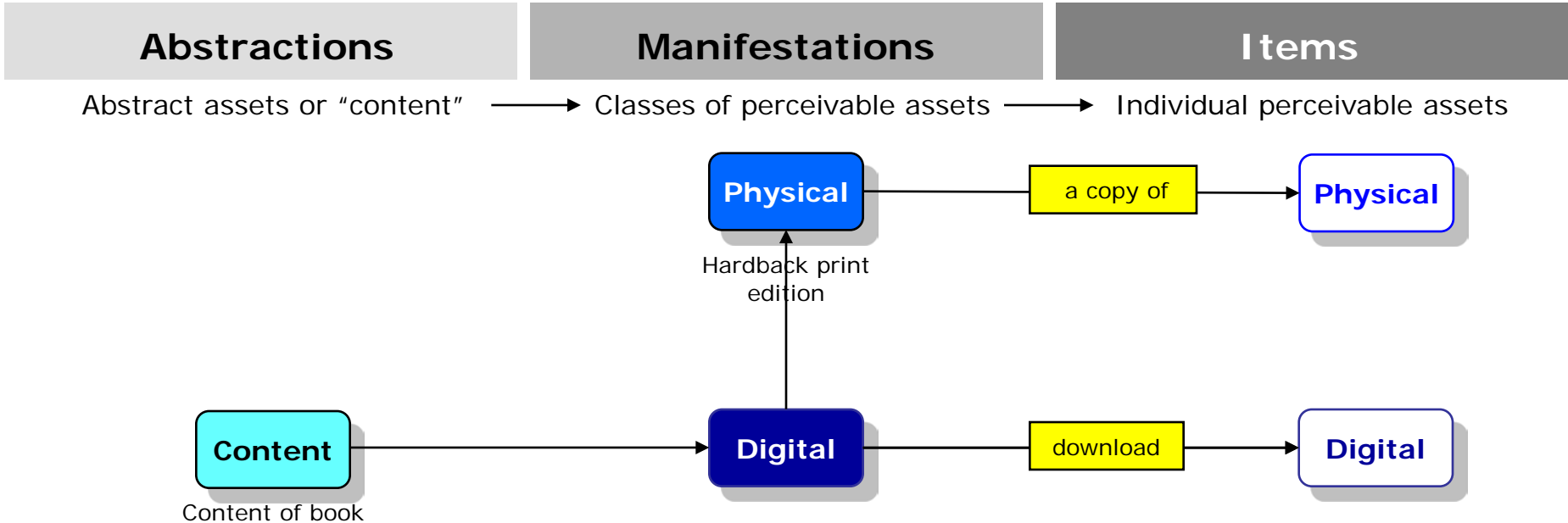
eg, an individual copy of a  
printed book  
music CD  
leaflet  
sculpture

**Digital**

eg, an individual copy of  
.pdf of an article  
.mp3 of a sound recording  
.exe of a program  
.gif of a photograph

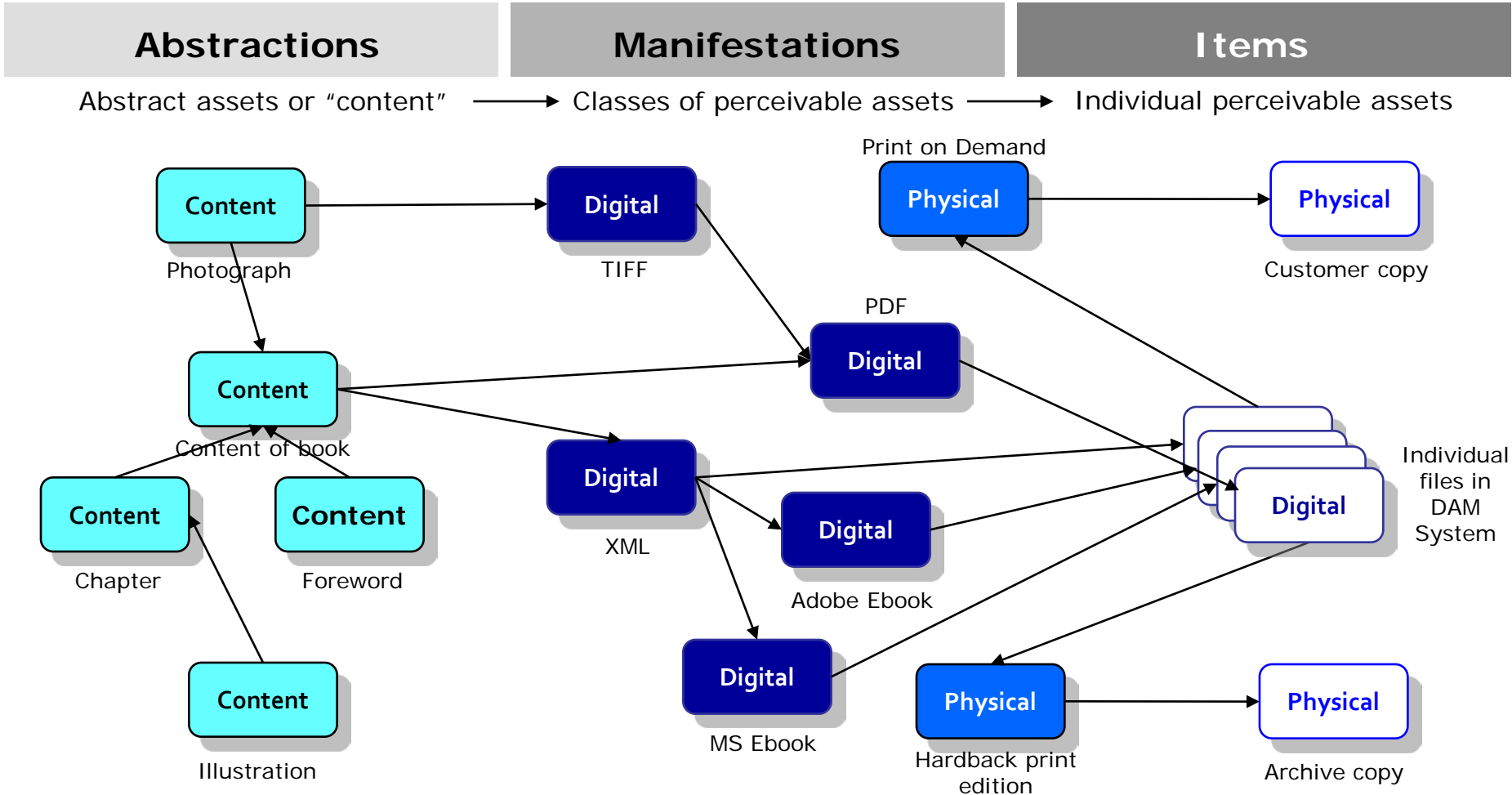
# Classes of asset

This simplified example shows just one asset and relationship of each type. In practice there may be multiple assets of all types with relationships at all levels, including between assets of the same type...



# Classes of publishers' assets

This is also **simplified**, but starts to show the complexity of identification and relationship that exists...



# Defining an “Asset”

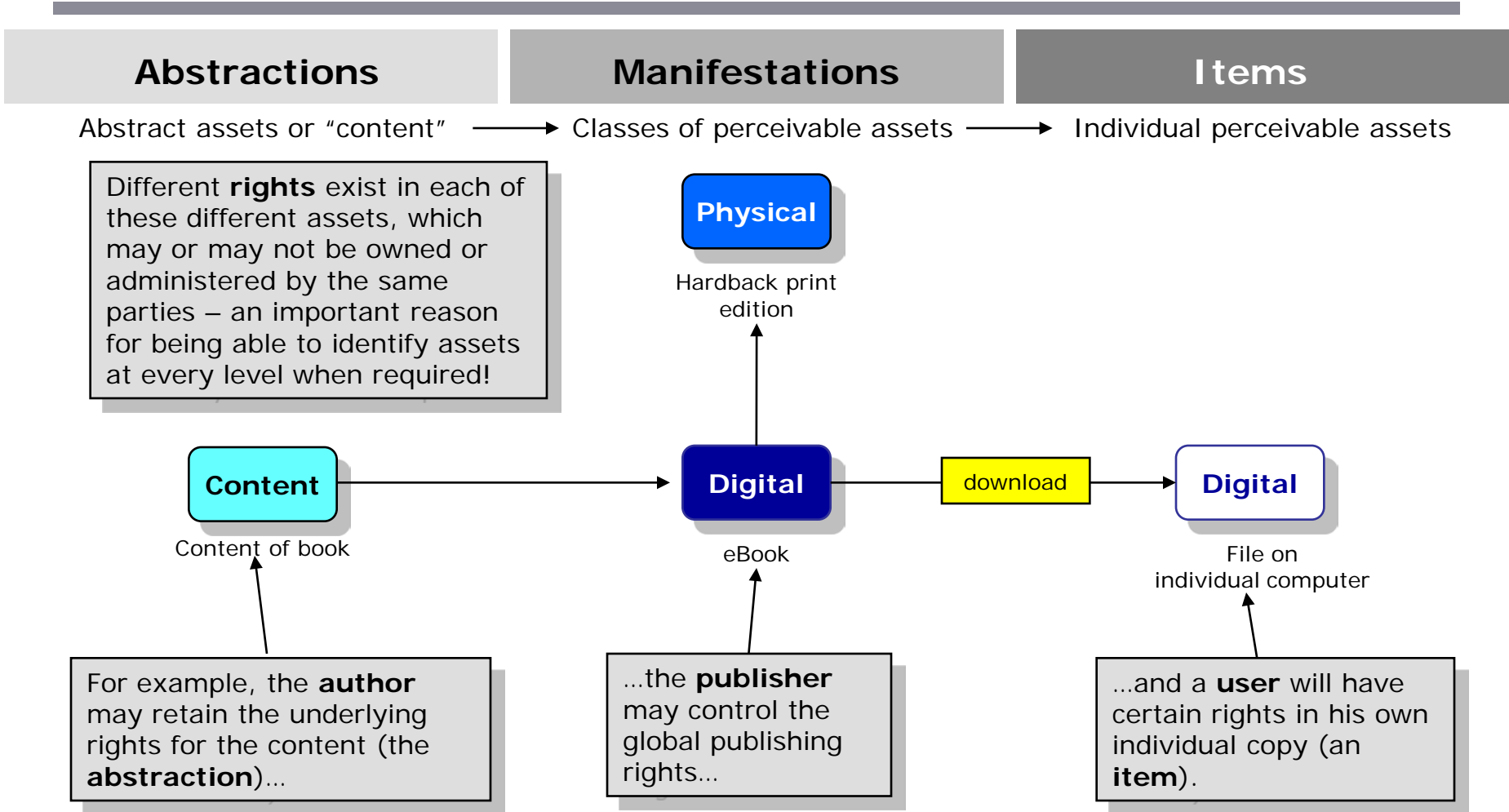
---

- ▶ So far we have talked about an Asset as if it were simply the equivalent of Content (or a Manifestation)
- ▶ But in reality, Content is an Asset only if you have the rights to use it

Content + Rights = Asset

Content – Rights = Liability

# Classes of publishers' assets



*The Digital Supply Chain: today and tomorrow:* London: 4 December 2008

# Assets and metadata

---

- ▶ Digital Asset Management isn't simply about putting things away – it's also about retrieving them for re-use
- ▶ Metadata is the key to effective DAM
  - ▶ Discovery – content metadata
    - > What is the asset? (eg the text of....; a picture of....)
  - ▶ Fitness for purpose
    - > Rights metadata: Can we re-use this asset in this context?
    - > Technical metadata: Is the file format appropriate?
- ▶ Identification is the critical item of metadata
  - ▶ “Are we talking about the same thing”?
  - ▶ The book industry recognised this very early – using the ISBN to identify products...
  - ▶ ...but then unfortunately went on to use it to identify everything else

# Assets and identity

---

- ▶ The “Master ISBN” is a commonly used as a proxy work identifier in publishers’ systems
  - ▶ The ISBN used to identify both a Work and a Product
  - ▶ Not a huge problem in the world of physical products...
  - ▶ ...but a growing one in the world of digital ones
  
- ▶ Granularity substantially adds to the challenge
  - ▶ Granularity of digital use....many different products (think of all those different ebook formats)
  - ▶ Granularity of digital content...many different items of content used in different contexts (eg the same content used in many different learning objects)

# The principle of functional granularity

---

- ▶ Why do we need identifiers?
  - ▶ Collocation – to identify when two things are “the same thing” *for our current purpose*
  - ▶ Disambiguation – to identify when two things are “different things” *for our current purpose*
- ▶ You can continue to identify things with the same identifier for as long as you are happy to treat them as “the same thing” – but when you need to distinguish between them, they need different identifiers.
- ▶ Using the same identifier for a Work and a Product creates particular challenges when it comes to communicating between systems

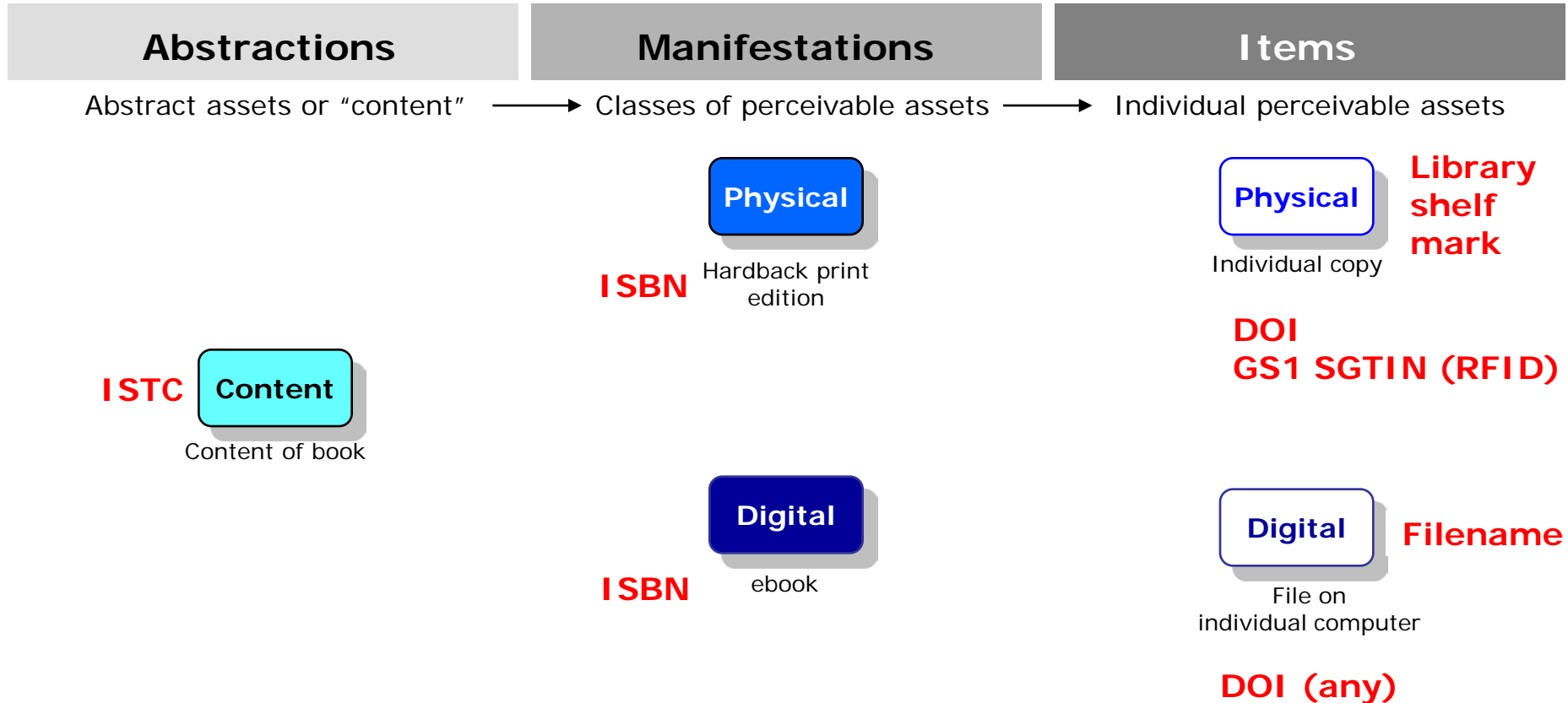
---

Clearly we need to be able to identify works (and many other things) unambiguously – but do we need anything other than internal proprietary identifiers?

# IDENTIFICATION STANDARDS

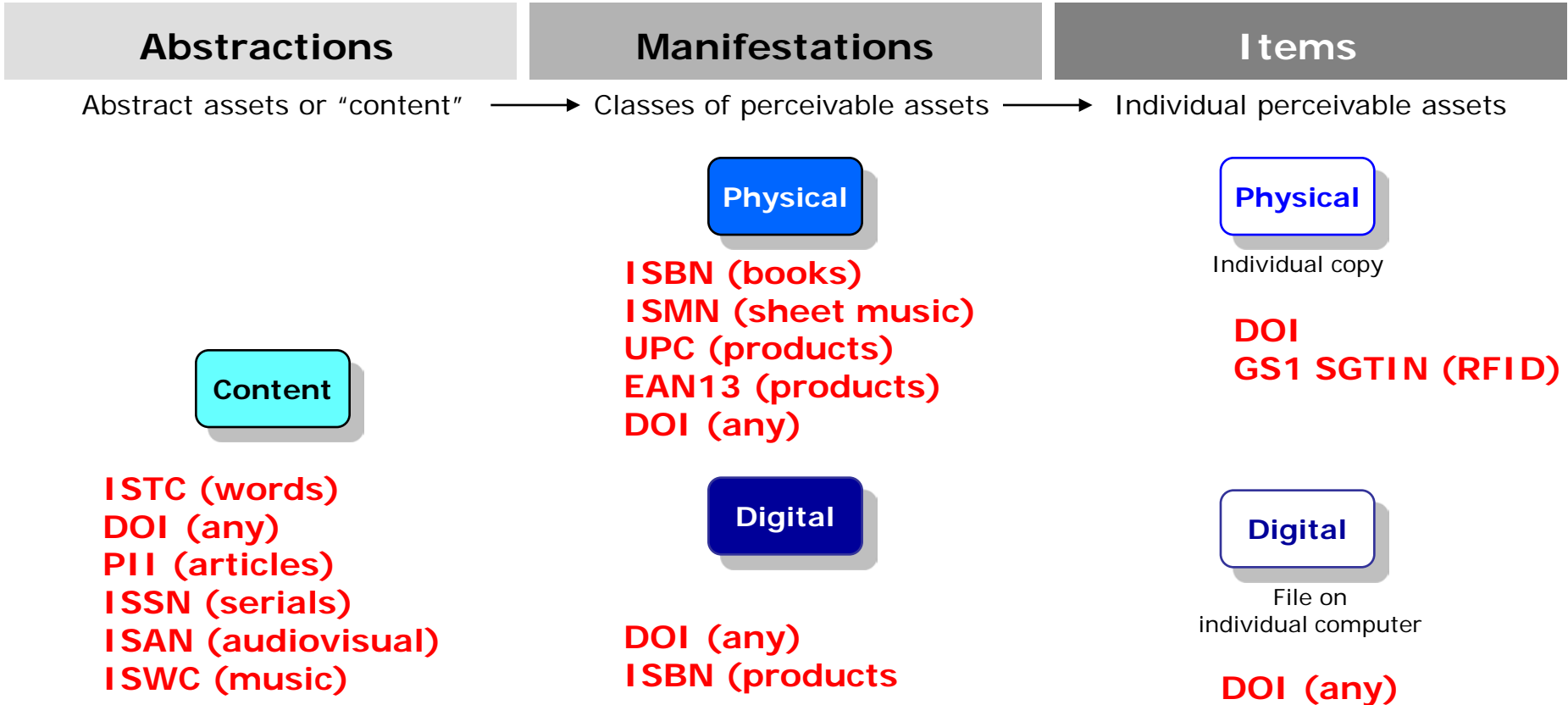
# Classes of publishers' assets

Examples of **standard identifiers** suitable for use for the assets we have already looked at



# Classes of asset

Some standard **identifiers** suitable for identifying assets of different types



# When do we need identification standards?

---

- ▶ Identifiers support communication between “systems”
- ▶ If all those systems are within the same organisation, then identifiers can be proprietary
  - ▶ Nobody (other than you!) cares much about how your DAM communicates with your distribution system or your royalty system
- ▶ As you need to communicate with other people’s systems, standard identifiers become increasingly helpful – particularly in supporting unambiguous many to many communication.
  - ▶ Common syntax
  - ▶ Common “identity model”
    - > What is being identified
    - > What are the granularity rules

# The challenge of a standard work identifier for books – different identity model requirements

---

Different users have different functional requirements – so different granularity:

- ▶ *Publishers* may need to collocate all the different manifestations of the same edition of a book
- ▶ *Retailers* may need to do the same...or may need to collocate all the different editions of “the same work” from different publishers
- ▶ *Librarians* may need to collocate the same work, but distinguish between what FRBR calls different “expressions” of “the same work”
- ▶ *Rights management organisations* may need to distinguish between different versions of “the same work” because of differences in rights ownership

# A soluble problem?

---

- ▶ Yes...
- ▶ ...but only at the expense of much more sophisticated metadata management ....
- ▶ ...(think of all those relationships)...
- ▶ ...and where is the cost benefit analysis...
- ▶ ...and we haven't even begun to talk through the problems of "fragments" of text...
- ▶ ...or photographs...

# The role of a work identifier in the publishing process

---

Thank you for your attention

[mark.bide@rightscom.com](mailto:mark.bide@rightscom.com)